

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/CN2006/001153

International filing date: 30 May 2006 (30.05.2006)

Document type: Certified copy of priority document

Document details: Country/Office: CN  
Number: 200510087707.1  
Filing date: 05 August 2005 (05.08.2005)

Date of receipt at the International Bureau: 12 September 2006 (12.09.2006)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

中华人民共和国国家知识产权局  
STATE INTELLECTUAL PROPERTY OFFICE  
OF THE PEOPLE'S REPUBLIC OF CHINA



## 证 明

本证明之附件是向本局提交的下列专利申请副本

申 请 日： 2005.08.05

申 请 号： 200510087707.1

申 请 类 别： 发明专利

发明创造名称： 一种端到端业务快速收敛的方法和路由设备

申 请 人： 华为技术有限公司

发明人或设计人： 熊怡、吴小前

中华人民共和国  
国家知识产权局局长

2006 年 9 月 6 日

## 权 利 要 求 书

- 1、一种端到端业务快速收敛的方法，其特征在于，包括：
  - a、在远端CE双归的PE中预先设置以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息；
  - b、检测隧道状态；
  - c、远端CE双归的PE根据检测的隧道状态信息、预先设置的所述隧道的转发路由信息进行业务转发。
- 2、如权利要求 1 所述的一种端到端业务快速收敛的方法，其特征在于，所述隧道为：包含内、外两层的隧道，且所述内层隧道为：VPN。
- 3、如权利要求 2 所述的一种端到端业务快速收敛的方法，其特征在于，所述步骤 a 具体包括：

远端 CE 双归的 PE 根据匹配策略将符合条件的优选、次优选的以远端 CE 双归的 PE 为起始节点、以与远端 CE 连接的 PE 为终止节点的转发路由信息设置在转发引擎的路由转发表中。
- 4、如权利要求 3 所述的一种端到端业务快速收敛的方法，其特征在于，所述步骤 a 具体包括：

将所述次优选的转发路由信息设置在路由转发表的优选的的转发表项中。
- 5、如权利要求 2 所述的一种端到端业务快速收敛的方法，其特征在于，所述步骤 b 具体包括：

远端 CE 双归的 PE 的控制层根据双向转发检测或隧道快速收敛在确定隧道的外层隧道状态发生变化时，更新其存储的隧道状态信息，并将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎。

6、如权利要求 5 所述的一种端到端业务快速收敛的方法，其特征在于，所述转发引擎的转发路由表中设置有隧道状态字段；

且所述步骤 b 中将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎的步骤具体包括：

远端 CE 双归的 PE 将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎的转发路由表中，并更新对应表项中隧道状态字段的内容。

7、如权利要求 5 或 6 所述的一种端到端业务快速收敛的方法，其特征在于，所述步骤 a 的至少两条隧道包括：互为备份的主用隧道和备用隧道；

且所述步骤 c 具体包括：

远端 CE 双归的 PE 在需要将业务通过主用隧道转发至所述远端 CE 时，获取并判断主用隧道状态信息；

如果主用隧道状态信息为可用，将所述业务通过主用隧道转发至所述远端 CE；

如果主用隧道状态信息为不可用，将所述业务通过备用隧道转发至所述远端 CE。

8、如权利要求 5 或 6 所述的一种端到端业务快速收敛的方法，其特征在于，所述步骤 a 中的至少两条隧道包括：互为负载分担的隧道；

且所述步骤 c 具体包括：

远端 CE 双归的 PE 在需要将业务通过互为负载分担的隧道转发至所述远端 CE 时，获取互为负载分担的隧道状态信息，并在互为负载分担隧道中的一条隧道不可用、其它隧道可用时，将所述业务通过可用的隧道转发至所述远端 CE。

9、如权利要求 5 或 6 所述的一种端到端业务快速收敛的方法，其特征在于，所述外层隧道不可用包括：隧道的终止节点异常。

10、一种路由设备，其特征在于，包括：存储模块、隧道状态检测模块和

转发模块;

存储模块: 存储有以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息和隧道状态信息;

隧道状态检测模块: 检测隧道状态, 并在隧道状态变化时, 更新存储模块中存储的隧道状态信息;

转发模块: 根据存储模块中存储的隧道状态信息、转发路由信息进行业务转发。

## 技术领域

本发明涉及网络通讯技术领域，具体涉及一种端到端业务快速收敛的方法和路由设备。

## 背景技术

目前，随着网络的高速发展，三网合一的需求日益迫切，运营商对网络故障时的业务收敛速度非常重视，在任何一个节点发生故障时，相邻节点业务倒换小于50ms，端到端业务收敛小于1s已经逐步成为承载网的门槛级指标。

为了达到相邻节点业务倒换小于50ms、端到端业务收敛小于1s的要求，MPLS TE FRR（多协议标签交换流量工程快速重路由）技术、IGP（域内网关协议）路由快速收敛技术等应运而生。

在CE（用户侧边界设备）双归PE（运营商边界设备）的网络模型下，MPLS TE FRR是目前解决网络故障时业务快速倒换最常用的技术之一，其基本实现原理为：在两个PE（运营商边界设备）设备之间建立端到端的TE隧道，并且为需要保护的主用LSP（标签交换路径）事先建立好备用LSP，当PE设备检测到主用LSP不可用如节点故障或者链路故障时，将流量倒换到备用LSP上，从而实现业务的快速倒换。

下面结合附图1对双归属网络模型下，基于MPLS TE FRR的业务快速倒换方法进行描述：

图1中，PE-E为远端CE双归的PE，PE-A和PE-B均为与远端CE连接的PE。设定用户侧边界设备CE-B访问用户侧边界设备CE-A的路径为：

CE-B—PE-E—P-C—PE-A—CE-A;

当PE-A节点故障之后，CE-B访问CE-A的路径收敛为：

CE-B—PE-E—P-D—PE-B—CE-A;

按照标准的MPLS L3 VPN（多协议标签交换三层虚拟私有网络）技术，首先，PE-A和PE-B都会向CE-A双归的PE-E发布指向CE-A的路由，并分配私网标签。

PE-E根据策略优选一个MP-BGP（Multi-Protocol Border Gateway Protocol，多协议承载边界网关协议）邻居发送的VPN V4路由，设定PE-E优选路由为PE-A发布的路由，则PE-E只将PE-A发布的路由信息如转发前缀、内层标签、选中的外层隧道等信息填写在转发引擎使用的转发项中，指导业务转发。

对于TE隧道起始节点PE-E和终止节点PE-A两个PE设备之间的链路故障和节点故障，MPLS TE FRR能够实现快速的业务倒换。

对于TE隧道起始节点PE-E和终止节点PE-A这两个PE设备的故障，在隧道起始节点故障时，通常采用的技术是CE-A通过双向路径检测等技术感知其直连的PE设备，如PE-A的故障，感知到故障之后，主动将流量发送给PE-B，以恢复业务。在隧道终止节点故障时，如PE-A节点故障时，PE-E只能通过BGP（边界网关协议）邻居DOWN或者外层LSP隧道不可用等信息感知到PE-A的故障，并重新优选PE-B发布的VPN V4路由，同时，将新的转发项下发到转发引擎的转发项中，从而完成端到端业务收敛。

在PE-E重新下发PE-B发布的路由到对应的转发项之前，由于PE-E转发引擎的转发项指向的外层LSP隧道的终点一直是PE-A，而PE-A节点故障，所以，在PE-A节点出现故障到PE-E重新下发PE-B发布的路由到对应的转发项之前这段时间内，CE-B是无法访问CE-A的，端到端业务中断。

在终止节点PE-B的PE设备故障时，业务恢复正常传输的时间主要取决于业务收敛的时间，而业务收敛时间与MPLS VPN内部路由的数量、承载网的跳数

10

密切相关，在典型组网中，业务收敛的时间一般在5s左右，无法达到端到端业务收敛小于1s的要求，并且端到端业务收敛时间还会随着MPLS VPN网络承载的私网路由数目的增加而显著增大。

因此，在CE双归PE的网络模型下，目前的MPLS TE FRR技术无法解决隧道隧道终止节点故障时，端到端业务快速收敛的问题。

## 发明内容

本发明的目的在于，提供一种端到端业务快速收敛的方法，通过检测隧道状态，并在隧道不可用时，利用预先设置的转发路由信息进行数据报文转发，有效提高了端到端业务的收敛速度，提高了业务可靠性。

为达到上述目的，本发明提供一种端到端业务快速收敛的方法，包括：

a、在远端CE双归的PE中预先设置以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息；

b、检测隧道状态；

c、远端CE双归的PE根据检测的隧道状态信息、预先设置的所述隧道的转发路由信息进行业务转发。

所述隧道为：包含内、外两层的隧道，且所述内层隧道为：VPN。

所述步骤a具体包括：

远端CE双归的PE根据匹配策略将符合条件的优选、次优选的以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的转发路由信息设置在转发引擎的路由转发表中。

所述步骤a具体包括：

将所述次优选的转发路由信息设置在路由转发表的优选的的转发表项中。

所述步骤b具体包括：

远端CE双归的PE的控制层根据双向转发检测或隧道快速收敛在确定隧道



的外层隧道状态发生变化时，更新其存储的隧道状态信息，并将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎。

所述转发引擎的转发路由表中设置有隧道状态字段；

且所述步骤 b 中将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎的步骤具体包括：

远端 CE 双归的 PE 将该隧道的外层隧道可用/不可用的状态信息下发至转发引擎的转发路由表中，并更新对应表项中隧道状态字段的内容。

所述步骤 a 的至少两条隧道包括：互为备份的主用隧道和备用隧道；

且所述步骤 c 具体包括：

远端 CE 双归的 PE 在需要将业务通过主用隧道转发至所述远端 CE 时，获取并判断主用隧道状态信息；

如果主用隧道状态信息为可用，将所述业务通过主用隧道转发至所述远端 CE；

如果主用隧道状态信息为不可用，将所述业务通过备用隧道转发至所述远端 CE。

所述步骤 a 中的至少两条隧道包括：互为负载分担的隧道；

且所述步骤 c 具体包括：

远端 CE 双归的 PE 在需要将业务通过互为负载分担的隧道转发至所述远端 CE 时，获取互为负载分担的隧道状态信息，并在互为负载分担隧道中的一条隧道不可用、其它隧道可用时，将所述业务通过可用的隧道转发至所述远端 CE。

所述外层隧道不可用包括：隧道的终止节点异常。

本发明还提供一种路由设备，包括：存储模块、隧道状态检测模块和转发模块；

存储模块：存储有以远端 CE 双归的 PE 为起始节点、以与远端 CE 连接的 PE 为终止节点的至少两条隧道的转发路由信息和隧道状态信息；

隧道状态检测模块：检测隧道状态，并在隧道状态变化时，更新存储模块中存储的隧道状态信息；

转发模块：根据存储模块中存储的隧道状态信息、转发路由信息进行业务转发。

通过上述技术方案的描述可知，通过在远端CE双归的PE中设置互为备份或互为负载分担的隧道的转发路由信息，并检测隧道的状态信息，使远端CE双归的PE能够在隧道不可用如隧道的终止节点出现异常时，直接根据备份隧道的转发路由信息进行业务转发，避免了重新优选路由的过程；通过采用BDF、隧道快速收敛等技术来感知隧道的不可用状态，使端到端故障感知时间小于500ms、甚至可以达到50ms，并且端到端故障感知时间与MPLS VPN网络承载的私网路由数目无关；通过将互为备份或互为负载分担的隧道的路由转发信息设置在一个转发表项中，使本发明能够快速、方便的获取备用隧道或负载分担隧道的路由转发信息；从而通过本发明提供的技术方案实现了提高端到端业务收敛速度，提高业务可靠性的目的。

#### 附图说明

图1是双归属网络模型示意图；

图2是本发明的路由设备示意图。

#### 具体实施方式

本发明的核心是：在远端CE双归的PE中预先设置以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息，检测隧道状态，远端CE双归的PE根据检测的隧道状态信息、预先设置的所述隧道的转发路由信息进行业务转发。

下面基于本发明的核心思想对本发明提供的技术方案做进一步的描述。

本发明中的隧道为包含内、外两层的隧道，其中，内层隧道可以为VPN，外层隧道可以为LSP隧道、GRE隧道（General Routing Encapsulation）、IPSec隧道（IP安全隧道）等。本发明不限制外层隧道的种类。

本发明首先需要选取以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息。远端CE双归的PE通过上述隧道均可以将近端CE的数据报文转发至远端CE。

当隧道的内层隧道为VPN、且外层隧道为LSP隧道时，至少两条转发路由信息的选取方法为：

远端CE双归的PE根据匹配策略选择符合预定条件的VPN V4路由，对于这些符合条件的VPN V4路由，本发明除了选取优选的路由信息，还需要选取次优选的路由信息。上述路由信息包括转发前缀、内层标签、选中的外层LSP隧道等。

然后，远端CE双归的PE应该存储上述优选的和次优选的路由信息。远端CE双归的PE可以将优选的路由信息、次优选的路由信息都填写在转发引擎的路由转发表的转发表项中。

上述以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道可以为互为备份的主用隧道和备用隧道，且备用隧道可以为一条或多条；也可以为互为负载分担的隧道，互为负载分担的隧道可以为两条或两条以上。

在设置了上述路由信息后，本发明需要检测上述隧道的状态，远端CE双归的PE可通过BFD（双向转发检测）、隧道快速收敛技术如LSP（标签交换路径）快速收敛等技术来检测隧道的状态，并在确定隧道的状态发生变化时，将其自身中设置的LSP隧道状态表中的对应标志设置为不可用，同时，将隧道状态信息下发至转发引擎，如远端CE双归的PE的控制层在确定主用隧道由可用状态转变为不可用状态时，在将其自身中设置的LSP隧道状态表中的对应隧道状态

标志修改为不可用的同时，将主用隧道的不可用状态信息下发到转发引擎，这样，远端CE双归的PE的转发引擎在需要将数据报文通过主用隧道进行数据报文转发时，如果确定主用隧道的状态为不可用时，则转发引擎可根据其存储的预先设置的备用隧道的转发路由信息进行数据报文转发。当然，远端CE双归的PE的转发引擎在确定主用隧道的状态为不可用时，还可以进一步判断备用隧道的状态，并在备用隧道为可用时，根据其存储的预先设置的备用隧道的转发路由信息进行数据报文转发。从而，实现了端到端业务的快速收敛。

当远端CE双归的PE的转发引擎需要将数据报文通过互为负载分担的隧道进行数据报文转发时，需要检测互为负载分担的各隧道的状态，并根据各隧道的状态来选取相应的隧道进行数据报文转发，如两条隧道互为负载分担，在其中一条隧道状态为不可用、且另一条隧道状态为可用时，将数据报文通过隧道状态为可用的隧道转发，其具体实现过程与上述通过主备用隧道数据报文转发的过程基本相同，在此不再详细描述。

为使远端CE双归的PE的转发引擎在进行数据报文转发时，能够方便的获取隧道状态信息，本发明可以在远端CE双归的PE的转发引擎的路由转发表中增设隧道状态字段，当转发表项中的路由转发信息为主、备用隧道的路由转发信息时，隧道状态字段可以仅标识主用隧道的状态信息，也可以标识出主、备用隧道的状态信息。当转发表项中的路由转发信息为互为负载分担隧道的路由转发信息时，隧道状态字段应标识出互为负载分担的各隧道的状态信息。这样，转发引擎在命中主用隧道或负载分担隧道的转发表项时，可以通过该表项中的隧道状态字段中的内容确定该隧道的状态。当然，隧道状态信息可以独立于路由转发表，单独存储。

为使远端CE双归的PE的转发引擎在进行数据报文转发、且主用隧道或互为负载分担隧道中的一条隧道故障时，能够方便的获取备用隧道或互为负载分担隧道中的其它隧道的路由转发信息，本发明可以将备用隧道的路由转发信息设

置在主用隧道的路由转发表项中，将互为负载分担隧道中的各隧道的路由转发信息设置在其中一条隧道的路由转发表项中，这样，转发引擎在命中主用隧道或一条负载分担隧道的转发表项、且该表项中主用隧道或其中一条负载分担隧道的状态为不可用时，可以直接从该表项中获取备用隧道或其它负载分担隧道的路由转发信息。

下面结合附图1、以主用隧道和备用隧道为例对本发明的端到端业务快速收敛的方法进行详细描述。

图1中，设定远端CE为CE-A，近端CE为CE-B，CE-A双归的PE为PE-E，与CE-A连接的PE为PE-A和PE-B。

PE-A和PE-B都会向CE-A双归的PE-E发布指向CE-A的路由，并分配私网标签。PE-E根据策略优选一个远端CE直连的PE设备发送的VPN路由，再次优选远端CE直连的另外一个VPN路由。

设定PE-E优选路由为PE-A发布的路由，次优选路由为PE-B发布的路由，则PE-E将PE-A、PE-B发布的路由信息如转发前缀、内层标签、选中的外层LSP隧道等信息填写在转发引擎使用的转发表项中，PE-A发布的路由为主用路由，PE-B发布的路由为备用路由。

PE-E将PE-A、PE-B发布的路由信息存储在转发引擎的转发表项中的方法为：将PE-A发布的优选路由信息存储在转发表项中，该表项中还应该有优选路由的隧道状态信息和次优选路由信息。

PE-E的控制层中设置有LSP隧道状态表，该表中存储有各隧道的状态信息。

在PE-A节点故障时，PE-E控制层通过BFD、LSP快速收敛等技术感知到PE-E与PE-A之间的外层隧道不可用，在典型组网中，端到端故障感知时间小于500ms，甚至可以达到50ms。

当PE-E的控制层感知到MPLS VPN依赖的外层LSP隧道不可用之后，将控制层的LSP隧道状态表中的对应标志设置为不可用，同时，将隧道不可用信息

下发到转发引擎中。

转发引擎命中一个路由转发项之后，如命中上述主用隧道的路由转发项，则检查该转发项对应的LSP隧道的状态，如果主用隧道状态为不可用，则使用本转发项中携带的次优路由的转发信息进行转发，这样，报文就会打上PE-B分配的内层标签，沿着PE-E与PE-B之间的外层LSP隧道交换到PE-B，再转发给CE-A，从而恢复CE-B到CE-A方向的业务，实现PE-A节点故障情况下的端到端业务的快速收敛。

本发明提供的路由设备如附图2所示。

在图2中，本发明的路由设备包括：存储模块、隧道状态检测模块和转发模块。

存储模块主用用于存储有以远端CE双归的PE为起始节点、以与远端CE连接的PE为终止节点的至少两条隧道的转发路由信息和隧道状态信息，上述转发路由信息和隧道状态信息可以以转发引擎路由转发表的形式存在，且上述至少两条隧道的转发路由信息和隧道状态信息可以存储于路由转发表的一个转发表项中，如主用隧道的路由信息中可以包括：主用隧道的路由信息、主用隧道的状态信息、备用隧道路由信息等。当然，上述路由信息和隧道状态信息也可以以其它方式存储。该存储模块可位于路由设备的各转发引擎中。

隧道状态检测模块主要用于检测隧道状态，并将隧道状态的变化信息下发至存储模块，更新存储模块中存储的隧道状态信息。隧道状态检测模块自身也可以存储隧道状态信息，该隧道状态信息可以以控制层的隧道状态表的形式存储。隧道状态检测模块可使用BFD、隧道快速收敛如LSP快速收敛等技术检测隧道状态。隧道状态检测模块可位于路由设备的控制层中，也可以位于各转发引擎中。

转发模块主要用于根据存储模块中存储的隧道状态信息、转发路由信息进行业务转发，如在命中存储模块中存储的一个路由转发表项时，如果该表项中

主用隧道状态信息为可用时，则根据该表项中主用隧道的路由信息进行业务转发，如果该表项中主用隧道状态信息为不可用时，则根据该表项中备用隧道的路由信息进行业务转发。转发模块可位于路由设备的各转发引擎中。

虽然通过实施例描绘了本发明，本领域普通技术人员知道，本发明有许多变形和变化而不脱离本发明的精神，本发明的申请文件的权利要求包括这些变形和变化。

18

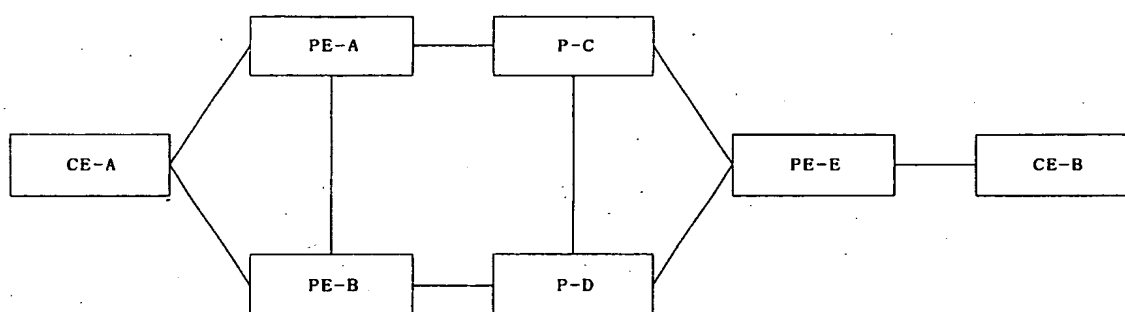


图1

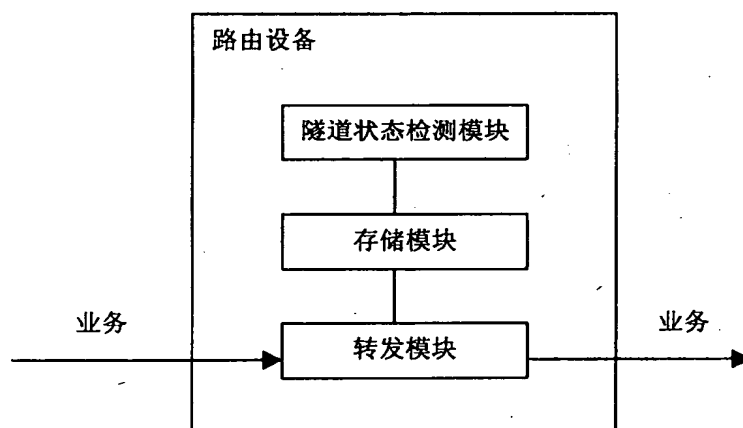


图2